

SEQUENCE LISTING

<110> Stomp, Anne-Marie
Dickey, Lynn
Gasdaska, John

<120> Expression of Biologically Active
Polypeptides in Duckweed

<130> 40989/237225

<150> US 60/293,330

<151> 2001-05-23

<150> US 60/221,705

<151> 2000-07-31

<160> 8

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 554

<212> DNA

<213> Zea mays

<400> 1

```

gatcaagtgc aaaggtccgc cttgtttctc ctctgtctct tgatctgact aatcttggtt 60
tatgattcgt tgagtaattt tggggaaagc ttcgtccaca gttttttttt cgatgaacag 120
tgccgcagtg gcgctgactt tgtatgctat cctgcaatcg tggatgaactt atgtctttta 180
tattccttcac taccatgaaa agactagtaa tctttctoga tgtaacatcg tccagcactg 240
ctattaccgt gtggtccatc cgacagtctg gctgaacaca tcatacgata ttgagcaaag 300
atctatcttc cctgttcttt aatgaaagac gtcattttca tcagtatgat ctaagaatgt 360
tgcaacttgc aaggaggcgt ttctttcttt gaatttaact aactcgttga gtggccctgt 420
ttctcggacg taaggccttt gctgctccac acatgtccat tcgaatttta cctgtgtttg 480
caagggcgaa aagtttgcat cttgatgatt tagcttgact atgcgattgc tttcctggac 540
cctgcagct gcgg

```

<210> 2

<211> 498

<212> DNA

<213> Artificial Sequence

<220>

<223> Duckweed codon optimized nucleotide sequence
encoding human alpha-2B interferon

<221> CDS

<222> (1)...(498)

<400> 2

```

tgc gac ctc ccc cag acc cac agc ctc ggg tcc cgc cgc acc ctc atg 48
Cys Asp Leu Pro Gln Thr His Ser Leu Gly Ser Arg Arg Thr Leu Met
1 5 10 15

```

ctg ctg gcg cag atg cgc cgc atc tcg ctc ttc agc tgc ctg aag gac	96
Leu Leu Ala Gln Met Arg Arg Ile Ser Leu Phe Ser Cys Leu Lys Asp	
20 25 30	
cgc cac gac ttc ggc ttc ccg cag gag gag ttc ggc aac cag ttc cag	144
Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Gly Asn Gln Phe Gln	
35 40 45	
aag gcc gag acg atc ccc gtg ctc cac gag atg atc cag cag atc ttc	192
Lys Ala Glu Thr Ile Pro Val Leu His Glu Met Ile Gln Gln Ile Phe	
50 55 60	
aac ctg ttc agc acc aag gac agc tcg gcc gcc tgg gac gag acc ctg	240
Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu Thr Leu	
65 70 75 80	
ctc gac aag ttc tac acc gag ctg tac cag cag ctc aac gac ctg gag	288
Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu Glu	
85 90 95	
gcg tgc gtg atc cag ggg gtt ggg gtt acg gag acg ccg ctg atg aag	336
Ala Cys Val Ile Gln Gly Val Gly Val Thr Glu Thr Pro Leu Met Lys	
100 105 110	
gag gac agc atc ctc gcc gtg cgc aag tac ttc cag cgc atc acg ctc	384
Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr Leu	
115 120 125	
tac ctc aag gag aag aag tac agc ccg tgc gcc tgg gag gtc gtt cgc	432
Tyr Leu Lys Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val Arg	
130 135 140	
gcc gag atc atg cgc tcc ttc agc ctg agc acc aac ctc cag gag agc	480
Ala Glu Ile Met Arg Ser Phe Ser Leu Ser Thr Asn Leu Gln Glu Ser	
145 150 155 160	
ctc cgc tcc aag gag taa	498
Leu Arg Ser Lys Glu *	
165	

<210> 3
 <211> 96
 <212> DNA
 <213> Oryza sativa

<400> 3
 accatgcagg tcttgaacac gatggtcaac aagcacttcc tctccctgtc cgtcctcatc 60
 gtcctcctcg ggctgagcag caacctcacc gccggc 96

<210> 4
 <211> 188
 <212> PRT
 <213> Homo sapiens

<400> 4
 Met Ala Leu Thr Phe Ala Leu Leu Val Ala Leu Leu Val Leu Ser Cys

1				5				10					15			
Lys	Ser	Ser	Cys	Ser	Val	Gly	Cys	Asp	Leu	Pro	Gln	Thr	His	Ser	Leu	
			20					25					30			
Gly	Ser	Arg	Arg	Thr	Leu	Met	Leu	Leu	Ala	Gln	Met	Arg	Arg	Ile	Ser	
		35					40					45				
Leu	Phe	Ser	Cys	Leu	Lys	Asp	Arg	His	Asp	Phe	Gly	Phe	Pro	Gln	Glu	
	50					55					60					
Glu	Phe	Gly	Asn	Gln	Phe	Gln	Lys	Ala	Glu	Thr	Ile	Pro	Val	Leu	His	
65					70					75					80	
Glu	Met	Ile	Gln	Gln	Ile	Phe	Asn	Leu	Phe	Ser	Thr	Lys	Asp	Ser	Ser	
				85					90					95		
Ala	Ala	Trp	Asp	Glu	Thr	Leu	Leu	Asp	Lys	Phe	Tyr	Thr	Glu	Leu	Tyr	
			100					105					110			
Gln	Gln	Leu	Asn	Asp	Leu	Glu	Ala	Cys	Val	Ile	Gln	Gly	Val	Gly	Val	
		115					120					125				
Thr	Glu	Thr	Pro	Leu	Met	Lys	Glu	Asp	Ser	Ile	Leu	Ala	Val	Arg	Lys	
		130				135					140					
Tyr	Phe	Gln	Arg	Ile	Thr	Leu	Tyr	Leu	Lys	Glu	Lys	Lys	Tyr	Ser	Pro	
145				150					155						160	
Cys	Ala	Trp	Glu	Val	Val	Arg	Ala	Glu	Ile	Met	Arg	Ser	Phe	Ser	Leu	
				165				170						175		
Ser	Thr	Asn	Leu	Gln	Glu	Ser	Leu	Arg	Ser	Lys	Glu					
			180					185								

<210> 5
 <211> 165
 <212> PRT
 <213> Homo sapiens

<400> 5																
Cys	Asp	Leu	Pro	Gln	Thr	His	Ser	Leu	Gly	Ser	Arg	Arg	Thr	Leu	Met	
1				5					10					15		
Leu	Leu	Ala	Gln	Met	Arg	Arg	Ile	Ser	Leu	Phe	Ser	Cys	Leu	Lys	Asp	
			20					25					30			
Arg	His	Asp	Phe	Gly	Phe	Pro	Gln	Glu	Glu	Phe	Gly	Asn	Gln	Phe	Gln	
		35					40					45				
Lys	Ala	Glu	Thr	Ile	Pro	Val	Leu	His	Glu	Met	Ile	Gln	Gln	Ile	Phe	
	50					55					60					
Asn	Leu	Phe	Ser	Thr	Lys	Asp	Ser	Ser	Ala	Ala	Trp	Asp	Glu	Thr	Leu	
65					70				75						80	
Leu	Asp	Lys	Phe	Tyr	Thr	Glu	Leu	Tyr	Gln	Leu	Asn	Asp	Leu	Met	Glu	
				85					90				95			
Ala	Cys	Val	Ile	Gln	Gly	Val	Gly	Val	Thr	Glu	Thr	Pro	Leu	Met	Lys	
			100					105					110			
Glu	Asp	Ser	Ile	Leu	Ala	Val	Arg	Lys	Tyr	Phe	Gln	Arg	Ile	Thr	Leu	
		115					120					125				
Tyr	Leu	Lys	Glu	Lys	Lys	Tyr	Ser	Pro	Cys	Ala	Trp	Glu	Val	Val	Arg	
	130					135					140					
Ala	Glu	Ile	Met	Arg	Ser	Phe	Ser	Leu	Ser	Thr	Asn	Leu	Gln	Glu	Ser	
145				150					155						160	
Leu	Arg	Ser	Lys	Glu												
				165												

<210> 6
 <211> 31

<212> PRT
<213> Oryza sativa

<400> 6
Met Gln Val Leu Asn Thr Met Val Asn Lys His Phe Leu Ser Leu Ser
1 5 10 15
Val Leu Ile Val Leu Leu Gly Leu Ser Ser Asn Leu Thr Ala Gly
20 25 30

<210> 7
<211> 31
<212> PRT
<213> Artificial Sequence

<220>
<223> Modified rice alpha-amylase signal peptide

<400> 7
Met Gln Val Leu Asn Thr Met Val Asn Lys His Phe Leu Ser Leu Ser
1 5 10 15
Val Leu Ile Val Leu Thr Val Leu Ser Ser Asn Leu Thr Ala Gly
20 25 30

<210> 8
<211> 21
<212> PRT
<213> Arabidopsis thaliana

<400> 8
Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser
1 5 10 15
Leu Ser Ser Ala Glu
20